



Identification and analysis of synthesis routes in complex catalytic reaction networks for biomass upgrading

Srinivas Rangarajan, Aditya Bhan*, Prodromos Daoutidis**

Department of Chemical Engineering & Materials Science, University of Minnesota, Minneapolis, MN 55455, United States

ARTICLE INFO

Article history:

Received 15 October 2012

Received in revised form

22 December 2012

Accepted 7 January 2013

Available online 8 February 2013

Keywords:

Network generation

Synthesis pathways identification

Group additivity based thermochemistry

Identification of synthesis routes to form fatty alcohols from biomass

QSPR-based property estimation of nonionic surfactants

ABSTRACT

In this paper, we present a method to construct and query the spectrum (>15,000) of potential biomass-derived compounds that can be synthesized using heterogeneous catalysis. Specifically, it involves the construction of an exhaustive network of reactions using RING, a rule-based network generator that we have developed, to identify potentially synthesizable compounds from biomass. Subsequently, we: (a) employ quantitative structure–property relationships to query compounds with desired properties, (b) seek different synthesis routes to these compounds, and (c) evaluate these routes in terms of stoichiometric, energetics, and physical parameters such as atom efficiency, enthalpy and free energy change of reactions, and aqueous–organic partition coefficients ($\log P$) of intermediates. Thermochemistry and $\log P$ are both estimated using group additivity methods, and all property estimations are performed on-the-fly during network generation. We present this method in context of identifying and evaluating heterogeneous catalytic routes from biomass to fatty alcohols that are potential constituents of nonionic surfactants. We also show that group additivity-based estimates of enthalpy of reaction are accurate to within 10 kJ/mol and $\log P$ values reliably predict the relative hydrophobicity of compounds. This method is generic, flexible, reliable and fast in terms of the scope of chemistry that can be considered, properties that can be included, predictive accuracy, and speed of execution; it can, therefore, be used to rapidly screen a large spectrum of compounds and synthesis routes in biomass conversion.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Biomass conversion to fuels and chemicals is a prospective “green” alternative to processing petroleum. Several catalytic routes have been proposed for converting biomass into valuable products [1,2]. This includes individual reaction systems such as acid-catalyzed fructose dehydration to 5-hydroxymethylfurfural (HMF) [3] and multi-step catalytic processes that convert biomass into fuels such as diesel [4,5]. Given the plethora of options for chemical transformations, a wide variety of chemicals can be synthesized from biomass [6–10]. Consequently, the network of all potential reactions to upgrade biomass – the synthesis network – is complex. The analysis of these complex networks – from identifying synthesis routes to calculating the energetics – requires a computational method having two features. First, the reaction network should be assembled in an automated manner. To this end, rule-based network generators that construct the reaction network

of a system exhaustively from fundamental reaction rules can be used. Second, the energetics of each reaction step in the network should be estimated rapidly. This is because, while *ab initio* computations have been used to calculate thermochemistry of synthesis routes [11], applying such a method to an entire complex network is computationally intractable.

In this paper, we explore the synthesis network of biomass conversion to fatty alcohols for the production of nonionic surfactants in terms of: (a) determining the spectrum of products that can be synthesized from biomass-derived precursors using a defined set of rules for acid, base, and metal catalysis, (b) selecting synthetically feasible compounds that have user-desired properties, (c) identifying the different routes to synthesize these desired products from biomass, and (d) comparing these routes in terms of parameters such as atom efficiency, selectivity, energetics, and two-phase separability. To address the computational challenge associated with the size of this problem, we employ: (a) a computational software we have developed – RING [12,13], a rule-based reaction network generation and analysis tool – to construct the network and identify synthesis routes, and (b) on-the-fly semi-empirical property prediction using group additivity for thermochemistry and quantitative structure–property relationships (QSPRs) for physical properties. Through this example, as a case study, we present

* Corresponding author.

** Corresponding author. Tel.: +1 612 625 8818; fax: +1 612 626 7246.

E-mail addresses: abhan@umn.edu (A. Bhan), daout001@umn.edu, daoutidi@cems.umn.edu (P. Daoutidis).

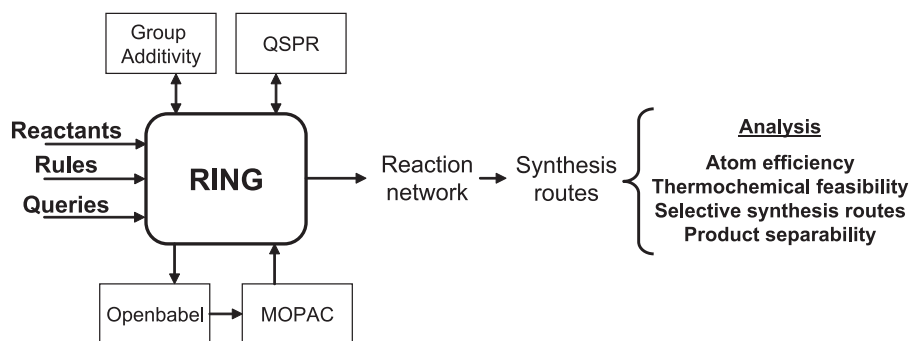


Fig. 1. A workflow of network generation, analysis, and semi-empirical molecular property relationship estimation with RING.

a scalable technique for rapidly screening the large spectrum of biomass-derived compounds and identifying and evaluating synthesis routes to valuable ones.

2. Computational methods

The reaction network was generated using RING [12,13], as discussed in Section 1. Inputs into RING include the set of initial reactants, the reaction rules describing the chemistries, and a set of post-processing network analysis features such as queries on molecules, reactions, pathways, and mechanisms in the network. A mechanism listing the steps that need to be taken (in a specific order) from the initial reactants to the final products, is, in effect, a synthesis route. The gas phase thermochemistry – enthalpy and entropy of formation, and heat capacity – of molecules at any specified temperature is calculated using group additivity. To this end, an additivity scheme has been implemented within RING that takes user-specified groups along with other inputs to calculate thermochemistry on-the-fly. For this work, groups covering hydrocarbons and oxygenates have been used [14–17]. Octanol-water partition coefficients ($\log P$) are calculated for potential biphasic reactions also using group additivity [18]. Quantitative structure–property relationships for nonionic surfactants (see Section 3.1), are also implemented in RING for on-the-fly physical property prediction.

One of the properties, surface tension (Section 3.1), requires the calculation of enthalpy of formation using PM3 [19] level of theory. To accommodate this calculation, we have integrated RING with the open source tool Openbabel [20] and MOPAC [21]. This allows for calculating enthalpy of formation of a generated molecule in a three-step procedure. First, when RING generates a new species, during the process of network generation, it supplies Openbabel with the SMILES [22] string of the species. Openbabel translates this SMILES string into a MOPAC-compatible input file with preliminary estimates of the 3D structure [20]. Second, MOPAC performs the geometry optimization and energy calculations to return an output file with the results. Third, RING parses the file, to obtain the relevant data that is required. This three-step procedure, albeit automated, was found to be the most time consuming step in property estimation. For the generated network, an analysis with a randomly selected set of 15 fatty alcohols showed that the mean error between calculating the thermochemistry using group additivity and PM3 level of theory was less than 2% (see Section S1 of supporting information for a detailed tabulation). The QSPR for surface tension was found to be insensitive to this level of error. Therefore, for initial querying analysis, we calculate the properties using group additivity. Subsequently, we refine the surface tension values using the enthalpy of formation value calculated at the PM3 level of theory (calculated using MOPAC) for those selected molecules. The accuracy of PM3 and conformational searching, as above, do not affect the estimation significantly (see Section S1

for more discussion). Fig. 1 provides a schematic of the work and information flow involved.

3. Fatty alcohols from biomass

3.1. Fatty alcohols

Fatty alcohols are alcohols with a large alkyl (C_8+) group. One of the major applications of fatty alcohols is in the synthesis of nonionic surfactants known as alcohol ethoxylates (AEs). These are compounds composed of one unit of a fatty alcohol and several units (up to 12 typically) of ethylene oxide (EO) [23] (Fig. 2). The alkyl group of the fatty alcohol constitutes the hydrophobic (or lipophilic) part, while the hydrophilicity is due to the EO units. Such surfactants find applications as constituents of detergents, cleaners, emulsifiers, etc. and are sold commercially [24]. The design of AEs involves tuning the size and structure of the hydrocarbon chain (for adjusting the hydrophobic properties), and increasing/decreasing the number of EO units. The desired physical properties that are typically targeted [25] include: (a) low critical micelle concentration (CMC) – the minimum concentration at which a surfactants aggregate to form a micelle, (b) high cloud point – the temperature above which the surfactant action ceases, (c) application-specific hydrophilic–lipophilic balance (HLB) – the balance between the hydrophobicity and lipophilicity of the molecule that determines if the surfactant is water or oil soluble, (d) low interfacial surface tension (γ) at a specified temperature and composition, (e) high biodegradation by microorganisms, measured in time taken to break the molecule down completely, and (f) aquatic toxicity – the concentration of surfactants in water above which it is lethal for aquatic life. Quantitative structure–property relationships have been developed to relate most of these properties with the structure of the surfactant [26–30]. Typically, these relationships involve calculating topological indices of a part of or the whole molecule or calculating energetic and thermochemical properties using semi-empirical Hartree–Fock theories such as PM3 and AM1, and then relating the desired physical properties to these parameters [25].

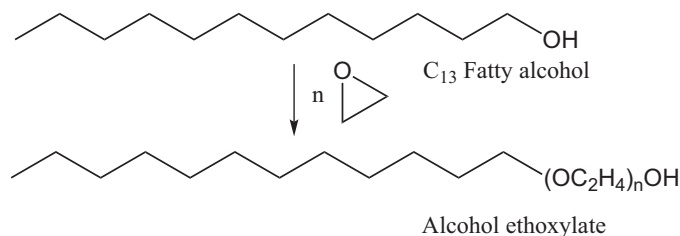


Fig. 2. A schematic of fatty alcohol and alcohol ethoxylates showing the hydrophobic (alkyl chain) and hydrophilic (ethylene oxide oligomeric chain) components.

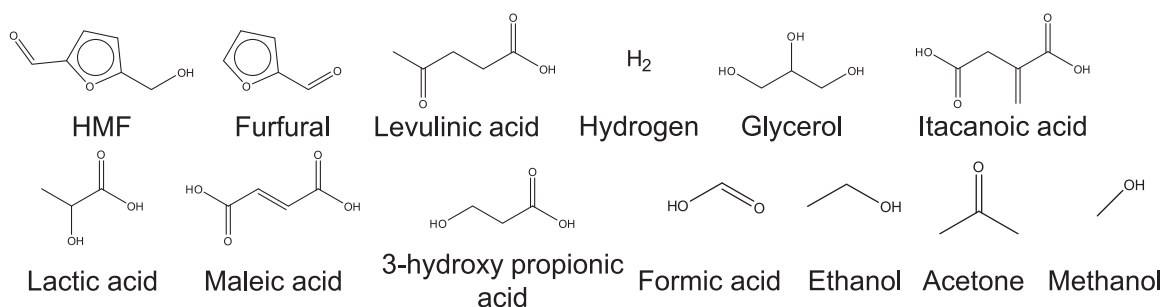


Fig. 3. Initial reactants input into RING. HMF stands for 5-hydroxymethylfurfural.

Being semi-empirical, these methods provide a fast and reasonably accurate [25] prediction of such properties.

3.2. Synthesis of desired fatty alcohols and alcohol ethoxylates

Several techniques have been conventionally used to synthesize fatty alcohols from vegetable oils [31], hydrocarbons [32], and small alcohols [33,34]. More recently, biochemical processes have been developed, by engineering microorganisms, to convert sugars into fatty alcohols [35,36]. Here, we explore heterogeneous catalytic routes to synthesize fatty alcohols from biomass-derived platform oxygenates.

The initial reactants are shown in Fig. 3, while the reaction rules that were used for generating the reaction network are listed in Table 1 along with representative examples of reactions. The reactants are taken from the US Department of Energy list of top 12 biomass-derived platform chemicals [6] and a subsequent study that revisited the topic [7]. The chemistries involve overall reaction rules comprising acid, base, and metal catalysis. Further, these rules comprise of C–C bond formation (aldol condensation, ketonization, Michael addition, alkylation), deoxygenation (dehydration, hydrogenolysis), and saturation (hydrogenation) – the three steps that lead to larger carbon chains, lower oxygen and higher hydrogen content in a molecule than the corresponding biomass-derived precursor. These reaction rules have been used based on their application in fine chemicals synthesis and biomass conversion [37–42]. The C–C bond formation steps were restricted to be applied only on molecules that have been generated in four or fewer steps from the initial reactants. Further, a global constraint was fixed to restrict the size of any generated molecule to be less than 15 atoms (excluding hydrogen). The network generated by RING has over 180,000 reactions and 50,000 species. To probe this large network, we now address the four questions raised earlier in context of the product and chemistry selection problem pertaining to the synthesis of fatty alcohols applicable as surfactants. Specifically, we present and discuss: (a) the spectrum of fatty monoalcohols formed in the network, (b) the subset of these alcohols that can lead to alcohol surfactants with properties close to commercially available ones, (c) the different atom efficient routes to synthesize these alcohols, and (d) the thermochemistry landscape and the biphasic separability aspects of these routes.

4. Results and discussion

4.1. Fatty alcohols product spectrum

In the large set of species in the generated network, we specifically identify saturated monoalcohols. Fig. 4 shows a spectrum of such alcohols. Clearly, a large number (16,000+ in total) and a variety (primary, secondary, linear, and branched) of monoalcohols exist in the network. This rich network is the starting point for identifying potential alcohol ethoxylates that can be synthesized.

4.2. Alcohol ethoxylates property estimation

The AEs are derived from the generated fatty alcohols, by successively adding ethylene oxide units to it in an automated manner within RING. Further, the surfactant properties are also estimated on-the-fly. Specifically, we consider AEs formed by adding 4–12 units of ethylene oxide per molecule of fatty alcohol. Thus, from each of the 16,000+ monoalcohols identified in the spectrum, we derive nine AEs. The properties of these surfactants are then estimated using the QSPRs discussed above (see Sections 2 and 3.1). The properties estimated include critical micelle concentration (CMC, in ppm), cloud point (CP, in $^{\circ}C$), hydrophilic–lipophilic balance (HLB, dimensionless), and surface tension at CMC (γ_{CMC} , in dynes/cm). Three commercially available surfactants, and their properties, are listed in Table 2 for reference. In the spectrum of product properties, we query for those surfactant compounds that fall in a target range of desired properties as determined from properties of commercial nonionic surfactants. The ranges were chosen in view of the error in the QSPRs, and of commercial surfactants being mixtures of several fatty alcohols. The results of the query for surfactants in the network in different target regions (I–III) around these three surfactants are presented in Table 3. Surface tension values in Table 2 are measurements at higher surfactant concentration (>0.1 wt%) than at the CMC; these values are, therefore, lower than γ_{CMC} . The queries take this into account by seeking a range higher in value than those reported in Table 2. Fig. 5 shows the structure of a few of the molecules identified in the three ranges. The structures of all molecules identified in each of the ranges are given in the supporting information (Section S2). One of the molecules in range I, lauryl alcohol (or 1-dodecanol), is a principal constituent of SURFONIC L12-8 [42,43], while the other structures identified in the search have different structures but have properties similar to that of lauryl alcohol. This combination of network generation

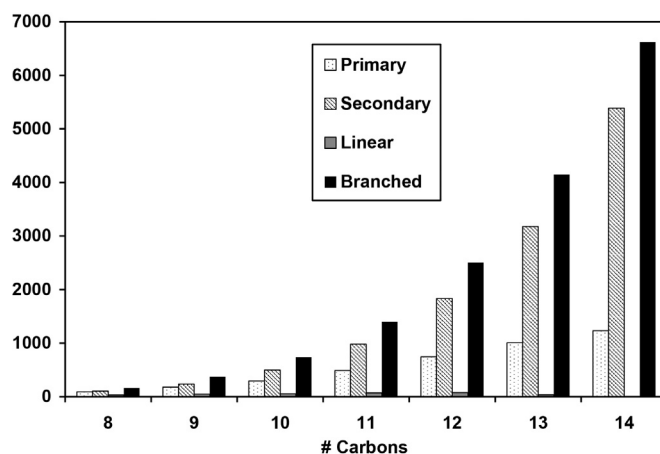

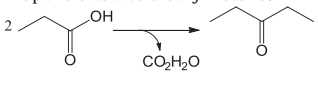
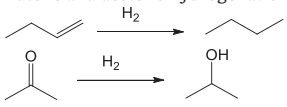
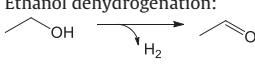
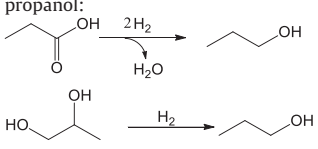
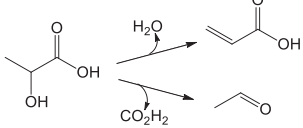
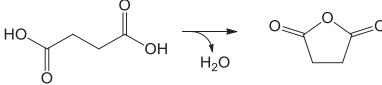
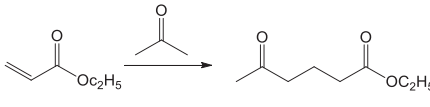
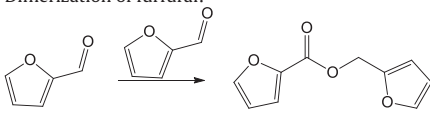
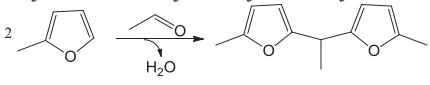
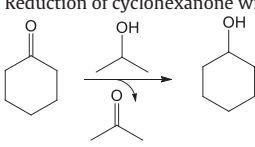


Fig. 4. The spectrum of monoalcohols generated in the network.

Table 1
Reaction rules and examples input into RING.

Reaction rules	Examples of reactions
Aldol condensation, with and without oxygen retention reversal	<p>Acetone to methyl isobutyl ketone (MIBK):</p> 
Ketonization of acids and esters	<p>Propanoic acid to diethyl ketone:</p> 
Hydrogenation of C=C and carbonyl groups, and aromatics	<p>Butene and acetone hydrogenation:</p> 
Dehydrogenation of alcohols	<p>Ethanol dehydrogenation:</p> 
Hydrogenolysis of acids, esters, cyclic ethers, and alcohols	<p>Propanoic acid to propanol, 1,3 propanediol to propanol:</p> 
Dehydration and decomposition of hydroxy acids	<p>Dehydration/decomposition of lactic acid:</p> 
Dehydration of diacids	<p>Succinic acid to succinic anhydride:</p> 
Michael addition	<p>Ethyl acrylate and acetone to form 5-oxohexanoic acid ethyl ester:</p> 
Tishchenko reaction	<p>Dimerization of furfural:</p> 
Alkylation and hydroxyalkylation of aromatics	<p>Alkylation of 2-methylfuran by acetaldehyde:</p> 
Meerwein Ponndorf Verley (MPV) reduction by isopropanol	<p>Reduction of cyclohexanone with isopropanol:</p> 

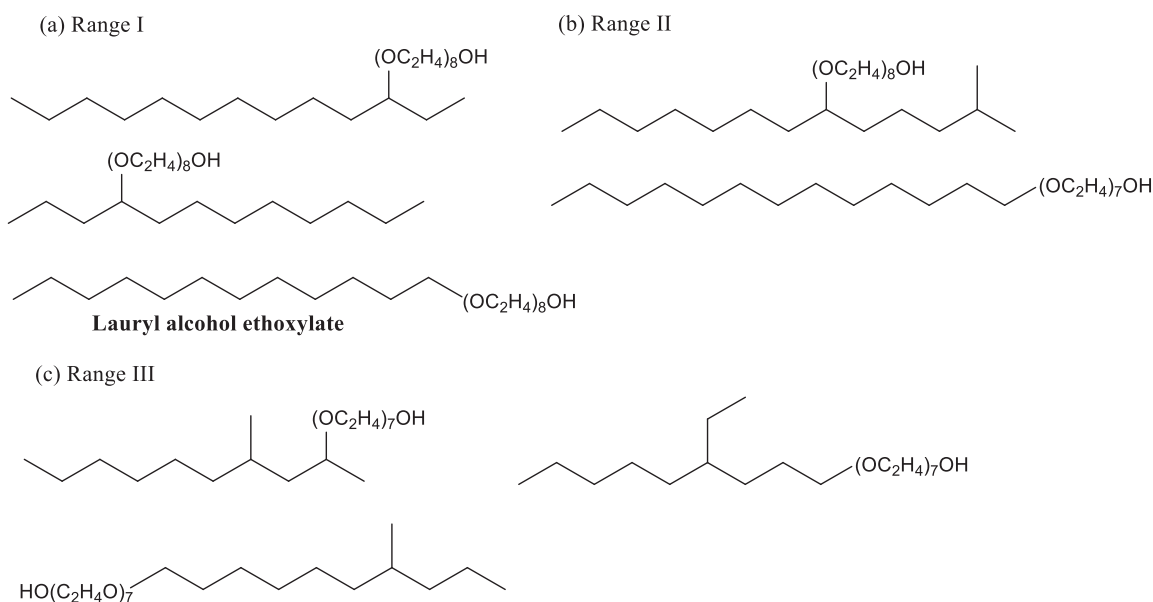


Fig. 5. Molecular structures of alcohol ethoxylates in each of the three ranges of queries.

Table 2
Commercial surfactants and their physical properties.

Name	CP (°C)	CMC (ppm)	HLB	γ (dynes/cm)
SURFONIC L12-8 [43,44]	80	77	13.6	29
SURFONIC L46-7 [43,44]	50	12	11.6	27
TERGITOL TMN-6 [45]	36	800	13.1	27

and QSPRs enables us to systematically identify potential surfactants that can be obtained from biomass through a combination of heterogeneous acid, base, and metal-catalyzed routes.

4.3. Synthesis routes to lauryl alcohol

Fig. 6 shows several routes to form lauryl alcohol. The synthesis route involving furfural, ethanol, and acetone (shown in bold arrows) represents an atom-efficient route for the synthesis of lauryl alcohol – atom efficiency described on the basis of the weight of useful product to that of byproducts. All these routes are similar in their sequence – C–C bond formation reactions (aldol condensation steps) occur first, followed by deoxygenation steps (hydrogenolysis). Further, dehydrogenation steps (ethanol to acetaldehyde) generate carbonyl groups for further C–C bond formation steps, while hydrogenation steps at the end of the sequence saturate the chains.

Each of the steps in Fig. 6 has a precedent in that a similar reaction has been reported in the context of chemical synthesis from oxygenates. Ethanol/methanol (and in general alcohol) dehydrogenation on metals is a well-established class of reactions used industrially [38,46], while the subsequent aldol condensation of aldehydes and ketones is similar to the condensation step in butanol and MIBK synthesis [47,48]. Lactic acid decomposition

has been noted on several heterogeneous catalysts as a major byproduct in lactic acid dehydration (with selectivities comparable to that of acrylic acid, the major product) [49,50]. The successive aldol condensation steps of furfural with acetaldehyde and of their condensed product with a five-carbon ketone (2-pentanone) are similar to the condensation steps involving furanic derivatives such as HMF with acetone or itself [4,5]. The subsequent hydrogenation of C=C, carbonyl groups, and aromatics have established precedents [4,5,38,40]. The hydrogenolysis steps to remove the hydroxy group in the hydroxyalkyl-substituted furan can be classified under the category of alcohol hydrogenolysis [38]. Finally, the conversion of alkyl-substituted tetrahydrofuran to a linear alcohol is analogous to selective hydrogenolysis of substituted cyclic ethers to obtain alcohols that has been recently reported [51].

Reactions in Fig. 6 do not take into consideration the selectivity of each step. Selectivity, however, is particularly important in this context because several condensation steps, including aldol condensation, can lead to undesired condensed products as byproducts. For example, Fig. 7 shows the possible condensation products of acetone and acetaldehyde; pent-3-en-2-one (marked 1 in Fig. 7) is the desired product. While including this cross condensation reaction, however, the possibility of forming the other three products cannot be ignored. One way to prevent such reactions from being included in the synthesis routes is by only allowing either self condensations or cross condensations wherein one of the co-reactants cannot condense with itself (for example, furfural and acetone, because the former does not have two hydrogen atoms α to the carbonyl group). Further, enforcing hydrogenation steps to occur successively will ensure that partial hydrogenation – another reaction step that could lead to multiple intermediates (and hence lower selectivity to the desired product) – is avoided.

Fig. 8 shows two selective synthesis routes to form lauryl alcohol (the difference is only in the manner 2-butenal is formed). These routes involve aldol condensation of furfural with acetone, and a second condensation with 2-butenal to form a C_{12} condensed molecule. The 2-butenal is, in turn, formed by aldol condensation of either acetone with formaldehyde or self-condensation of acetaldehyde. The C_{12} molecule is subsequently hydrogenated and undergoes hydrogenolysis to form the fatty alcohol. This synthesis route contrasts with that depicted in Fig. 6 wherein the products of furfural-acetaldehyde and acetaldehyde-acetone aldol

Table 3
Queried surfactant ranges and the number of identified molecules.

Query	CP (°C)	CMC (ppm)	HLB	γ (dynes/cm)	Number of surfactants identified
Range I	50–100	0–100	12–14	27–35	12
Range II	30–60	0–30	11–13	25–35	8
Range III	30–40	500–1000	12.5–13	27–34	18

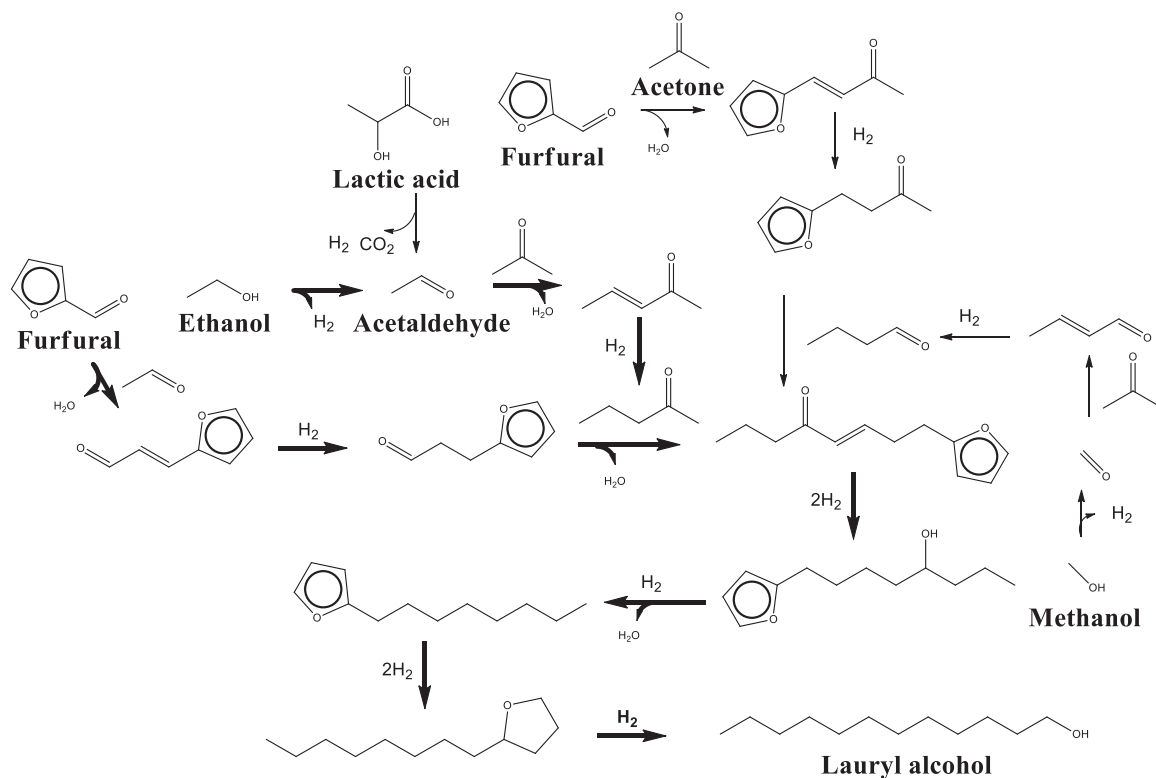


Fig. 6. Lauryl alcohol synthesis routes from biomass-derived oxygenates. An atom-efficient route is shown with bold arrows.

condensation were hydrogenated to saturate the C–C double bonds. Postponing this intermediate hydrogenation step in the routes as shown in Fig. 8 ensures that intermediates such as 2-butenal and formaldehyde cannot undergo aldol condensation further. The atom efficiency of these routes (shown in Fig. 8) matches the one in bold arrows in Fig. 6. This sequence is indeed the preferred choice in the synthesis of diesel-range alkanes from oxygenates derived from C₅ and C₆ sugars. For example, as shown in Fig. 9, Huber et al. [5] suggest a route involving multiple aldol condensation steps of 5-hydroxymethylfurfural (HMF) followed by hydrogenation/hydrogenolysis to form long chain alkanes. West et al. [52] also discuss the synthesis of alkanes of different molecular weights wherein HMF-Acetone-HMF condensation is carried out prior to hydrogenation. Thus, plausible synthesis routes to lauryl alcohols can be systematically identified and screened on the basis of prior chemistry knowledge and/or stoichiometric parameters such as atom efficiency.

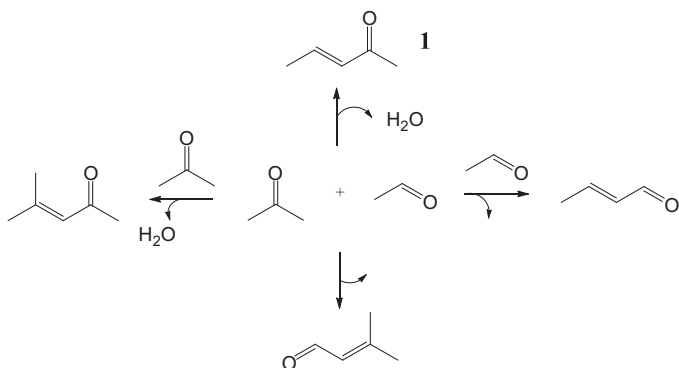


Fig. 7. Possible products of acetone-acetaldehyde aldol condensation including the desired product Pent-3-en-2-one (marked 1).

4.4. Thermochemical analysis of synthesis pathways

The analysis in the previous section demonstrates that while multiple synthesis routes may have identical overall stoichiometry, the sequence and the choice of chemistry steps can lead to considerable differences in other parameters such as yield/selectivity.

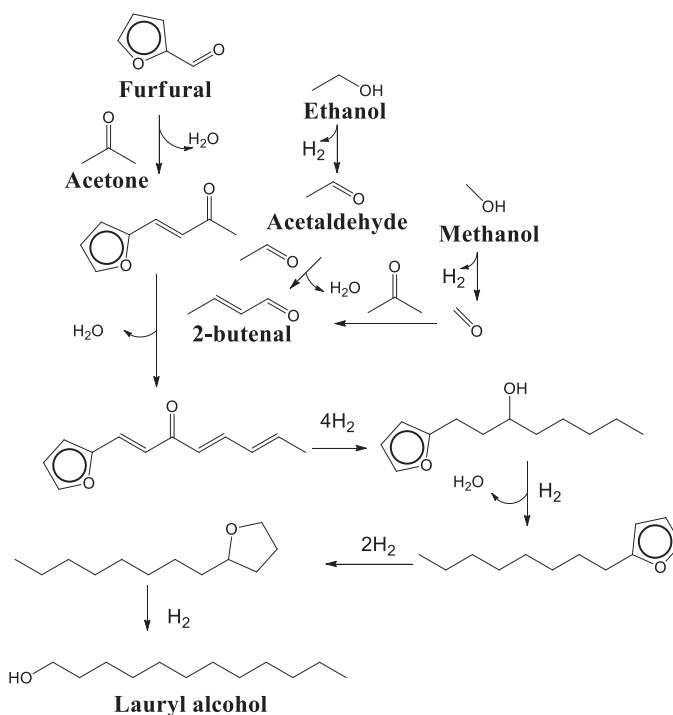


Fig. 8. Selective synthesis of lauryl alcohol from furfural, ethanol, methanol, and acetone.

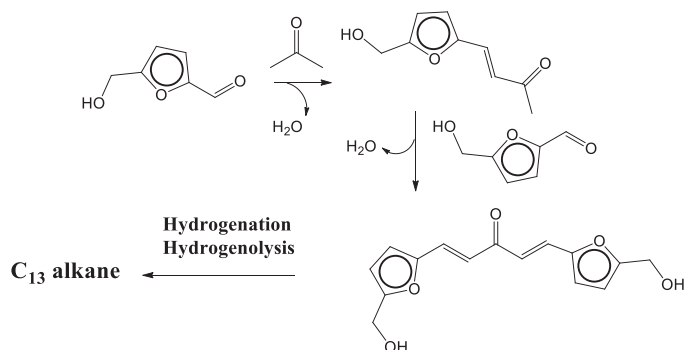


Fig. 9. Synthesis scheme for converting biomass-derived oxygenates to diesel. Adapted from Ref. [5].

In this section, these routes are compared in terms of different thermochemical parameters. As discussed in Section 2, RING calculates the thermochemistry on-the-fly using group additivity. While this method has shown very good accuracy for hydrocarbons (<1 kcal/mol [15]), the errors could be larger for oxygenates encountered in biomass conversion. To test the accuracy of group additivity, a sample set of 30 reactions and more than 50 compounds was considered. The species considered include hydrocarbons and oxygenates such as butane, hexane, benzene, lactic acid, pentanol, dimethyl ether, propane diol, acetaldehyde, furfural, tetrahydrofuran, etc., while reactions considered involve these compounds. The enthalpy of formation (and reaction enthalpy change) at 298 K and 1 atm pressure was estimated using RING and compared with values reported in the literature (predominantly from NIST webbook). The sample set and the estimations are given in the supporting information (Section S3).

Table 4 gives the error statistics for species enthalpy of formation and reaction enthalpy change. It can be noted that the mean unsigned error is remarkably small (5.2 kJ/mol). The large standard deviation indicates that while mean unsigned error can be less than 1 kcal/mol, specific compounds can have larger deviations. For example, group additivity value for enthalpy of formation of 5-hydroxymethylfurfural (HMF) is ~26 kJ/mol (~6 kcal/mol) more negative than the value reported in the literature by Assary et al. [11]. This can be attributed to the limited data set used to formulate the group additivity that does not accurately capture the effects of substitution in furans. The unavailability of accurate experimental thermochemistry data prohibits further scrutiny of group additivity methods for large oxygenates such as those seen in Figs. 6 and 8.

Given the error in calculation of species enthalpy of formation, it would be expected that the errors would propagate further for the estimation of reaction enthalpy change, thereby leading to much larger deviations. However, as seen from the results shown in Table 4, the error statistics of reaction enthalpy change appear to be identical to that of species enthalpy of formation for both unsigned and absolute errors. The P-value in both cases is large enough implying that the null hypothesis – the two distributions are statistically similar in nature – holds to a level of significance

much greater than the typical value of 5% [53]. This statistical equivalence can be attributed to the nature of group additivity. The error in each constituent group of a molecule contributes to the total deviation in calculating species enthalpy of formation. On the other hand, only those groups that participate in a reaction contribute to the deviation in the reaction enthalpy; the contribution of groups preserved in the reaction gets canceled out in the calculation. This leads to an average deviation in reaction enthalpy change comparable to that of species enthalpy.

Examples of this “cancellation” effect are shown in Table 5 wherein G4MP2 (from Assary et al. [11]) and group additivity values for some compounds and reactions are compared. It is clear that while the estimated enthalpy of formation of compound A in the table (condensate of HMF and acetone) is close to that predicted using G4MP2, estimations of the other three compounds deviate by 8–20 kJ/mol. On the other hand, the deviation in the estimation of enthalpy change of the two hydrogenation reactions shown in Table 5 is smaller than that observed for compounds. This is because group additive enthalpy of formation values are more negative than their corresponding G4MP2 values in every case; thus, the errors cancel out to an extent while calculating the reaction enthalpy. Therefore, although the enthalpy of formation of the C₁₂ condensate in Fig. 8 may be estimated using group additivity with significant error, the hydrogenation reaction enthalpy of that compound can be calculated with reasonable accuracy (≤10 kJ/mol) because the furanic group and the nature of substitution (C substitution in position 2) is preserved in the reaction.

In general, the signed error follows a normal distribution with zero mean for the training set used to obtain the additivity values [54]. In the present case, the additivity values have been obtained from multiple sources and our thermochemistry dataset is different from that used in those studies. The slight deviation of the mean signed error from zero for both the species and reactions enthalpy values reflects these differences. Furthermore, the standard deviation of signed/unsigned error of reaction enthalpy is larger than that of the species enthalpy. This observation is consistent with the results of Sutton and Vlachos [54] and is evident on the basis of statistical arguments. The sum of ‘n’ normally distributed random variables also follows normal distribution with a mean (and variance) equal to the sum of the mean (and variance) of those variables. Consequently, the variance of the sum is larger than that of the individual variables. Since reaction enthalpy is a sum of enthalpy of species involved in the reaction (taking into consideration the appropriate stoichiometric coefficient), the variance of errors in the reaction enthalpy is larger than that of the species enthalpy. Indeed, it must follow that the variance of species enthalpy is larger than the variance of individual group additivity values.

Fig. 10 shows the thermochemistry of the reaction steps in Fig. 8 and of one of the routes in Fig. 6. The enthalpy (and free energy) change of each reaction step is given at 1 atm pressure and temperature 600 K for each species. The temperature has been chosen to be considerably greater than the boiling point of lauryl alcohol (532 K [55]) to compare the gas phase energetics of all reactions at the same physical conditions.

It can be noted that while dehydrogenation of ethanol is endothermic, but feasible (negative ΔG_{rxn}), the subsequent aldol condensation steps with acetaldehyde or acetone are both mildly endothermic and non-spontaneous. These reactions require higher temperature or partial pressure of the reactants for thermochemical feasibility. The same is also true for aldol condensation of furfural with acetone/acetaldehyde. Assuming, as a first approximation, that these molecules behave as an ideal gas at 1 atmosphere and 600 K, the equilibrium conversion is approximately 35% and 10% for acetaldehyde-acetone and furfural-acetaldehyde condensation, respectively. The aldol condensation of compound marked 1 in Fig. 10 (Furfural-acetone condensation product) with

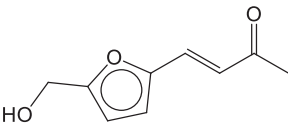
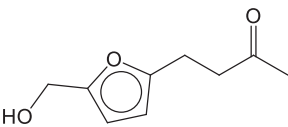
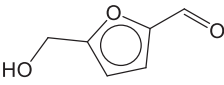
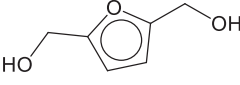
Table 4
Error statistics for species and reaction enthalpy estimation using group additivity.

	Signed error ^a (kJ/mol)		Absolute error (kJ/mol)	
	Mean	Std. deviation	Mean	Std. deviation
Species	−1.40	9.36	5.21	7.86
Reactions	1.59	11.28	5.36	9.32
P-value ^b		0.23		0.94

^a Errors are calculated as (group additivity value – experimental value).

^b P-value is the two-sided value calculated with t-test assuming normal distribution, unequal variances and sample size.

Table 5
Comparisons between G4MP2 [11] and group additivity estimations.

Species/reactions	Enthalpy (ΔH_f or ΔH_{rxn}) in kJ/mol		
	Group additivity	G4MP2 [11]	Difference
 A	–332.67	–331.06	–1.61
 B	–444.38	–436.39	–7.99
 HMF	–360.02	–339.42	–20.6
 DHMF	–411.47	–401.28	–10.19
A + H ₂ → B	–111.71	–105.34	–6.37
HMF + H ₂ → DHMF	–51.46	–61.86	10.41

2-butanone is exothermic and thermochemically spontaneous, while that of compound 2 (furfural-acetone condensation product) with 2-butanone is mildly endothermic and equilibrium limited (equilibrium conversion of 24%). It can be noted that subsequent hydrogenation of all the condensed products is both exothermic and spontaneous; this could possibly drive an equilibrium-limited aldol condensation step. For example, Kunkes et al. [56] showed that 2-hexanone condensation/hydrogenation over Pd/CeZrO_x at high pressures (5–26 bar) in the gas phase leads to the C₁₂ condensed product in high selectivity (>80%). Thermochemistry calculations (see [56]) showed that aldol condensation is equilibrium-limited at such conditions, while hydrogenation steps are irreversible. Thus, the authors concluded that the thermodynamically downhill hydrogenation step drives the equilibrium limited condensation step.

The C₁₂ condensates (3 and 4 in Fig. 10) are further hydrogenated and undergo hydrogenolysis in successive exothermic and spontaneous overall reactions. It is, however, important to note that hydrogenation of the carbonyl group (subsequent to saturating the C–C double bonds in the successive hydrogenation steps of compound 3) is not spontaneous at the conditions of 1 atm and 600 K ($\Delta H_{rxn} = -57.15$ kJ/mol, $\Delta G_{rxn} = 24.65$ kJ/mol). However, being exothermic, the equilibrium is shifted in the forward direction and at 400 K, ΔG_{rxn} is negative (–2.33 kJ/mol).

Hydro-dearomatization of compound 5 (Fig. 10) to form the corresponding tetrahydrofuran derivative is also exothermic and equilibrium-limited with a ΔG_{rxn} equal to 14.22 kJ/mol (see Fig. 10). Further, there is a net difference in the moles of product formed

compared to the reactants. Thus, the equilibrium of this reaction is shifted by changes in pressure and temperature. At 400 K (and 1 atm), ΔG_{rxn} is considerably negative (–38.55 kJ/mol). Alternatively, at 600 K and 10 atmosphere total pressure (90% hydrogen partial pressure), calculations, assuming ideality of gases, give an 80% equilibrium conversion.

The results discussed above outlining various routes for the synthesis of lauryl alcohols show that it is possible to rapidly (a) determine exothermicity/endothermicity of reactions in several synthesis routes, (b) identify thermodynamic bottlenecks (equilibrium-limited reactions), (c) postulate if some reactions can potentially drive others, and (d) predict physical conditions that could drive a reaction in the preferred direction.

4.5. Biphasic separability analysis of synthesis pathways

Several biomass conversion processes are carried out in liquid phase. In such cases, the product formed can be separated from the reactants by partitioning into a second phase within a reactor (biphasic reactors) [3,57] or in a downstream solvent separation unit [52]. Specifically, biphasic systems can drive the equilibrium forward and prevent the products from reacting further. For example, dehydration of fructose is carried out in a biphasic reactor to separate the product HMF from the reaction phase (aqueous phase containing acid) into an organic phase [3]. We have implemented on-the-fly prediction of octanol-water partition coefficients using group additivity in RING (see Section 2) and below, we discuss the use of this descriptor to assess biphasic separability.

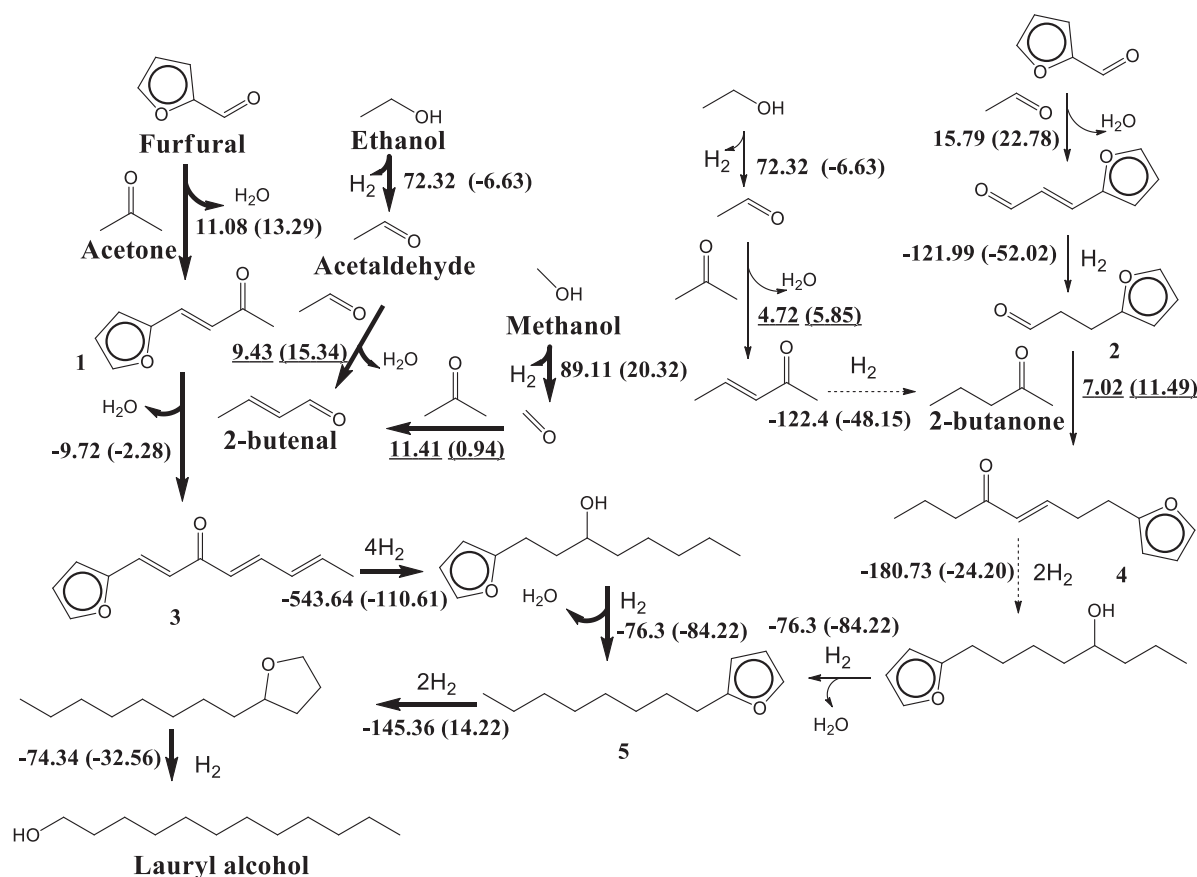


Fig. 10. Gas phase thermochemistry analysis of synthesis routes to lauryl alcohol. The route with bold arrows is a selective route without cross condensation steps involving molecules with α -hydrogens with respect to the carbonyl group. Enthalpy (and free energy) change (kJ/mol) of each reaction at 1 atm and 600 K is given. Thermochemistry values of equilibrium-limited reactions are underlined and hydrogenation steps that can potentially drive a preceding equilibrium-limited reaction are marked by dashed arrows.

Biphasic separability can, in general, be inferred from the partition coefficient between the aqueous phase and the organic phase. Ideally, partition coefficients between the actual organic phase (which could be a mixture of many solvents) and aqueous phase need to be evaluated; however, octanol–water partition coefficients (reported in logarithmic units, $\text{Log } P$) can give an approximate measure of the hydrophobicity/hydrophilicity of the molecule. Positive (or negative) values imply hydrophobicity (or hydrophilicity). A second advantage of using the $\text{Log } P$ measure is that these values can also be calculated using group additive techniques [18,58]. In this section, we analyze the synthesis routes to lauryl alcohol in terms of separability of reactants and products.

Fig. 11 shows the $\text{Log } P$ value of initial reactants, intermediates, and lauryl alcohol for the synthesis routes shown in Fig. 10. Table 6

compares experimental and estimated $\text{Log } P$ values for several oxygenates. Also given in the table are estimates from another additive method (CLOGP [59]) accessible through commercial software. Clearly, the method implemented in RING predicts the magnitude of hydrophobicity correctly in most cases, though the prediction is incorrect for acetone. The deviations, however, are consistent with the reported standard deviation of about 0.6 Log units [18]. CLOGP predictions are better; however it incorrectly predicts acetaldehyde to be hydrophilic. For compounds 1 and 4 in Fig. 10, both of which are unsaturated carbonyl derivatives of furan, the differences between the two methods are less than 0.3 units. Further, relative hydrophilicity (or hydrophobicity) values are predicted correctly by the $\text{Log } P$ estimation method. For example, glycerol, having a higher degree of hydrogen bonding per molecule than ethanol is expected, and observed, to be more hydrophilic than the latter. This is captured correctly in the $\text{Log } P$ values estimated by RING. Further, for a given class of molecules such as alkanes, alcohols, and aldehydes, it is observed that increasing carbon number results in a higher hydrophobicity [59]. This relative increase is also seen for $\text{Log } P$ estimates; for example, compound 4 is estimated to be more hydrophobic than compound 1 owing to its larger size. Thus, $\text{Log } P$ values assessed using RING are a reliable estimate of the octanol–water partition coefficient value.

From Fig. 11, it can be noted that several reactions have reactants and products differing in their solubility. Specifically, furfural–acetone (or acetaldehyde) condensation to form compound 1 (or compound 6), condensation of 1 with 2-butenal to form compound 3, condensation of compound 2 with 2-butanone to form compound 4, and the two hydrogenolysis reactions forming compound

Table 6
Experimental and estimated $\text{Log } P$ for a representative set of oxygenates.

Compounds	Octanol–water partition coefficient (in $\text{Log } P$ units)		
	Experimental	Estimated [18]	CLOGP [62]
Ethanol [60]	−0.3	−0.003	−0.24
Acetaldehyde [60]	0.45	0.20	−0.22
Furfural [61]	0.46	0.97	0.67
Dodecanol [60]	5.13	3.90	5.05
Glycerol [61]	−1.76	−1.43	−1.54
Acetone [61]	−0.24	0.59	−0.22
Lactic acid [61]	−0.72	−0.87	−0.73
Compound 1 ^a	–	1.37	1.32
Compound 4 ^a	–	3.01	2.69

^a in Fig. 11.

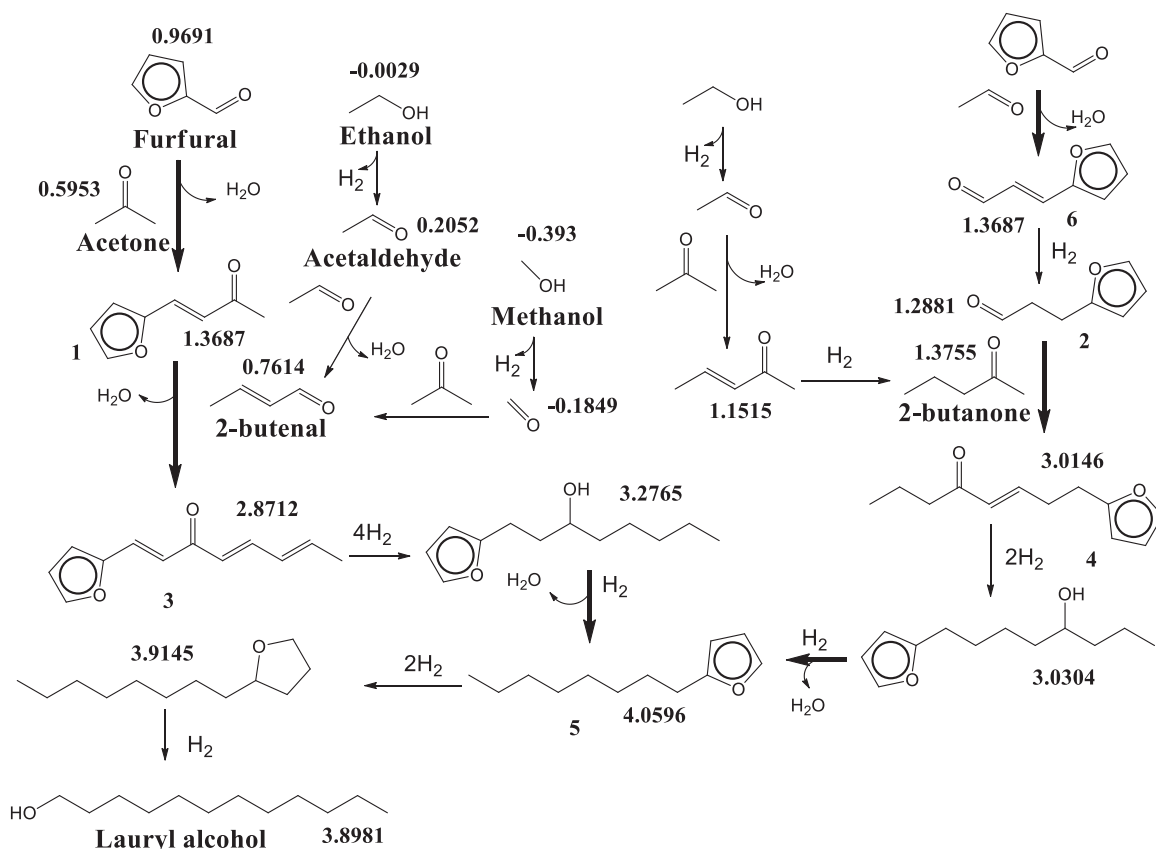


Fig. 11. Octanol-water partition coefficients of species in the synthesis routes to lauryl alcohol. The values corresponding to each species are calculated as the logarithm of the equilibrium partition coefficient value. Reactions marked with bold arrows can potentially be carried out in a biphasic aqueous-organic system.

5 have products that are at least 0.5 units more hydrophobic (a factor of three more soluble in octanol) than the reactants. These reactions, therefore, are potential biphasic systems. Indeed, several reports in the literature specifically employ biphasic systems (reactors/separators) for condensation/hydrogenolysis reactions for upgrading biomass. West et al. [52] discuss a process wherein the condensation and the final hydrogenation steps are followed by solvent-separation units where the aqueous contents are recycled while the contents of the organic phase are separated and either fed to the next stage or drawn off as products. Zapata et al. [57] use hybrid catalysts composed of base or metal catalysts fused to carbon nanotubes in an emulsion to simultaneously carry out condensation/hydrogenation reactions and separate the products. On-the-fly estimation of Log *P*, thus, enables identifying potential candidates wherein differences in solubility can be exploited for enhancing product selectivity and/or purification.

5. Discussion

We describe a computational method that combines automated network generation with semi-empirical molecular property prediction to screen large numbers of compounds and synthesis routes entailed in biomass conversion. In the preceding section, an application in the context of synthesizing fatty alcohols is discussed; however, the method adopted is not restricted to fatty alcohol synthesis alone and is generic enough to be applied to other classes of chemistries and compounds. For example, biomass conversion to fuels such as gasoline and diesel can also be explored in a similar manner, exploiting the vast resource base of structure–property relationships that exists for calculating properties such as octane

rating, cetane index, flash points, density, viscosity, cold flow properties, and boiling points [63–66].

There are three distinct advantages of this method: speed and automation, scalability, and reliability. Network generation, synthesis routes identification, and property-based selection/screening of molecules/routes are all performed by RING quickly and in an automated manner. The network represents the synthetically feasible set of molecules from initial reactants; the size of the network (in terms of the number of species and reactions), therefore, is a fractional subspace of the total chemical spectrum of possibilities. This implies that the computational cost of property estimation and identification of synthesis routes is significantly smaller than if the entire chemical spectrum had to be evaluated. Further, molecular property prediction methods are quantitatively and/or qualitatively reliable, as discussed in detail in Section 4.

At this stage of screening, the kinetics of each reaction step is not considered. That each reaction is a potential complex network consisting of several types of reactive intermediates and reactions depending upon the chemistry is also neglected. For example, aldol condensation is represented as an overall reaction step; but this reaction, however, in itself has several elementary step reactions that ultimately lead to the formation of the condensate. Therefore, though cross-condensation reactions are prevented while identifying selective routes in Section 4.3, several by-products of base catalysis are neglected. Furthermore, the thermochemical analysis discussed above only involves estimating gas phase properties, and consequently liquid phase energetics is not captured. To account for some of these, other semi-empirical correlations can be included. Kinetics can be used to evaluate synthesis routes by including rate parameters derived using linear free-energy relationships

involving energetic properties of reactants/products, such as in alcohol dehydration [67,68], or by using rate rules formulated on the basis of regression from experimental data [69,70], if available for all the specified chemistries. Liquid phase thermochemistry can be estimated using semi-empirical methods [71,72]. These methods have been proven to be fast and reliable in comparison to more sophisticated ab initio calculations [73]. The method we describe above is flexible in accommodating any type of semi-empirical correlation as a means of assessing properties and energetics of reactions or molecules.

6. Conclusion

We have combined network generation and post-processing pathway querying analysis and semi-empirical molecular property estimation in RING to enumerate and assess the synthetically feasible spectrum of fatty alcohols obtained from catalytic biomass conversion. Specifically, we identify those synthetically feasible fatty alcohol compounds that possess properties that make them desirable constituents of nonionic surfactants and identify and evaluate synthetic routes to form these compounds. Generation of the reaction network resulted in a spectrum of more than 16,000 monoalcohols. A query of the network for compounds having properties in a desirable range around that of a commercially available surfactant identified several choices, one of which was lauryl alcohol – a constituent alcohol of that surfactant. We systematically enumerate synthesis routes to lauryl alcohol from biomass-derived platform compounds and compared these routes on several parameters such as atom efficiency, thermochemistry, and aqueous-organic separability. We observed that gas phase thermochemistry predicted using group additivity is a reliable estimate of the species and reaction enthalpy with mean (and maximum) absolute deviations for a test set composed of hydrocarbons and oxygenates being ~ 5 kJ/mol (and ~ 20 kJ/mol). We also observed that, in several cases, the deviation of group additive estimates from ab initio computational chemistry predictions for reaction enthalpy was significantly smaller (by 5 – 10 kJ/mol) than that for species enthalpy of formation. Octanol-water partition coefficients estimated using group additive techniques predicted, in most cases, the hydrophobicity/hydrophilicity of biomass-derived oxygenates and the relative aqueous-organic solubilities of reactants and compounds in reactions correctly. We posit that this method of concurrent network generation and property estimation is generic and flexible enough to be applied to several classes of chemistries and compounds.

Acknowledgments

The authors would like to thank Chintamani Siddeshwar, Department of Computer Science and Engineering, University of Minnesota for help with developing the framework for interconnecting Openbabel, MOPAC, and RING. Financial support from the Initiative for Renewable Energy (Large Grant: RL-0004-09) at the University of Minnesota, the National Science Foundation Emerging Frontiers in Research and Innovation program, grant # 0937706, DTI award from the Digital Technology Center at the University of Minnesota, and The Dow Chemical Company is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.apcatb.2013.01.030>.

References

- [1] G.W. Huber, S. Iborra, A. Corma, *Chemical Reviews* 106 (2006) 4044–4098.
- [2] A. Corma, S. Iborra, A. Velty, *Chemical Reviews* 107 (2007) 2411–2502.
- [3] Y. Roman-Leshkov, J.N. Chheda, J.A. Dumesic, *Science* 312 (2006) 1933–1937.
- [4] E.L. Kunkes, D.A. Simonetti, R.M. West, J.C. Serrano-Ruiz, C.A. Gartner, J.A. Dumesic, *Science* 322 (2008) 417–421.
- [5] G.W. Huber, J.N. Chheda, C.J. Barrett, J.A. Dumesic, *Science* 308 (2005) 1446–1450.
- [6] Top Value Added Chemicals From Biomass. Vol. 1 – Results of Screening for Potential Candidates from Sugars and Synthesis Gas, U.S. Department of Energy, Washington, DC, 2004.
- [7] J.J. Bozell, G.R. Petersen, *Green Chemistry* 12 (2010) 539–554.
- [8] F. Cherubini, A.H. Strømman, *Biofuels, Bioproducts, and Biorefining* 5 (2011) 548–561.
- [9] J. van Haveren, E.L. Scott, J. Sanders, *Biofuels, Bioproducts, and Biorefining* 2 (2007) 41–57.
- [10] P. Gallezot, *Chemical Society Reviews* 41 (2012) 1538–1558.
- [11] R.S. Assary, P.C. Redfern, J.R. Hammond, J. Greeley, L.A. Curtiss, *Chemical Physics Letters* 497 (2010) 123–128.
- [12] S. Rangarajan, A. Bhan, P. Daoutidis, *Industrial & Engineering Chemistry Research* 49 (2010) 10459–10470.
- [13] S. Rangarajan, A. Bhan, P. Daoutidis, *Computers & Chemical Engineering* 45 (2012) 114–123.
- [14] S.W. Benson, *Thermochemical Kinetics*, 2nd ed., John Wiley & Sons, Inc, New York, 1976.
- [15] M.K. Sabbe, M. Saeys, M.-F. Reyniers, G. Marin, V. Van Speybroeck, M. Waroquier, *Journal of Physical Chemistry A* 109 (2005) 7466–7480.
- [16] M.K. Sabbe, F. De Vleeschouwer, M.-F. Reyniers, M. Waroquier, G.B. Marin, *Journal of Physical Chemistry A* 112 (2008) 12235–12251.
- [17] S.S. Khan, X. Yu, J.R. Wade, R.D. Malmgren, L.J. Broadbelt, *Journal of Physical Chemistry A* 113 (2009) 5176–5194.
- [18] S.A. Wildman, G.M. Crippen, *Journal of Chemical Information and Computer Science* 39 (1999) 868–873.
- [19] J.J.P. Stewart, *Journal of Computational Chemistry* 10 (1989) 209–220.
- [20] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G. Hutchinson, *Journal of Cheminformatics* 3 (2011) 33.
- [21] MOPAC, <http://openmopac.net/> (accessed 20.08.12).
- [22] D. Weininger, *Journal of Chemical Information and Computer Science* 28 (1988) 31–36.
- [23] K. Holmberg, B. Jönsson, B. Kronberg, B. Lindman, *Introduction to Surfactants, in Surfactants and Polymers in Aqueous Solution*, 2nd ed., John Wiley & Sons Ltd, Chichester, 2003 (Chapter 1).
- [24] C.L. Edwards, in: N.M. van Os (Ed.), *Nonionic Surfactants: Organic Chemistry*, CRC Press, New York, 1997, p. 111.
- [25] J. Hu, X. Zhang, Z. Wang, *International Journal of Molecular Sciences* 11 (2010) 1020–1047.
- [26] P.D.T. Huibers, V.S. Lobanov, A.R. Katritzky, D.O. Shah, M. Karelson, *Langmuir* 12 (1996) 1462–1470.
- [27] P.D.T. Huibers, D.O. Shah, A.R. Katritzky, *Journal of Colloid and Interface Science* 193 (1997) 132–136.
- [28] M.-L. Chen, Z.-W. Wang, H.-J. Duan, *Journal of Dispersion Science and Technology* 30 (2009) 1481–1485.
- [29] Z.-W. Wang, J.-L. Feng, H.-J. Wang, Z.-G. Cui, *Journal of Dispersion Science and Technology* 26 (2005) 441–447.
- [30] D.W. Roberts, *Journal of Surfactants and Detergents* 3 (2000) 309–315.
- [31] U.R. Kreutzer, *Journal of the American Oil Chemist's Society* 61 (1984) 343–348.
- [32] E. Billig, D.R. Bryant, *Oxo Process Kirk-Othmer Encyclopedia of Chemical Technology*, Wiley Interscience, New Jersey, 2000.
- [33] G. Knothe, *Lipid Technology* 14 (2002) 101–104.
- [34] A.J. O'Lenick, *Journal of Surfactants and Detergents* 4 (2001) 311–315.
- [35] E.J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McLure, S.B. Del Cardayre, J.D. Keasling, *Nature* 463 (2010) 559–562.
- [36] CODEXIS, <http://www.codexis.com/chemicals> (accessed 20.08.12).
- [37] H. Hattori, *Chemical Reviews* 95 (1995) 537–558.
- [38] R.A. Sheldon, H. van Bekkum, *Fine Chemicals through Heterogeneous catalysis*, Wiley-VCH Verlag GmbH, Weinheim, Germany, 2001.
- [39] U. Meyer, H. Gorzawski, W.F. Holderich, *Catalysis Letters* 59 (1999) 201–206.
- [40] A. Corma, O. de la Torre, M. Renz, N. Vollandier, *Angewandte Chemie-International Edition* 50 (2011) 2375–2378.
- [41] I.T. Horvath (Ed.), *Encyclopedia of Catalysis*, Wiley-VCH, Weinheim, 2002.
- [42] C.A. Gartner, J.C. Serrano-Ruiz, D.J. Braden, J.A. Dumesic, *ChemSusChem* 2 (2009) 1121–1124.
- [43] SURFONIC® L12-8 Technical Bulletin Huntsman Corporation; Huntsman Metalworking Chemicals – Product Information Brochure.
- [44] *Handbook of Green Chemicals*, Synapse Information Resources Inc, Endicott, NY, 2004.
- [45] DOW Tergitol TMN-6 (90% Surfactant) Technical Data Sheet, <http://www.dow.com/surfactants/products/branched.htm> (accessed 20.08.12).
- [46] H.F. Rase, *Handbook of Commercial Catalyst: Heterogeneous Catalysts*, CRC Press, Boca Raton, Florida, 2000 (Chapter 8).
- [47] J.I. Di Cosimo, C.R. Apesteguia, M.J.L. Gines, E. Iglesia, *Journal of Catalysis* 190 (2000) 261–275.
- [48] M.J. Climent, A. Corma, S. Iborra, *Chemical Reviews* 111 (2011) 1072–1133.

- [49] J.C. Serrano-Ruiz, J.A. Dumesic, *Green Chemistry* 11 (2009) 1101–1104.
- [50] G.C. Gunter, R.H. Langford, J.E. Jackson, D.J. Miller, *Industrial and Engineering Chemistry Research* 34 (1995) 974–980.
- [51] M. Chia, Y.J. Pagaín-Torres, D. Hibbitts, Q. Tan, H.N. Pham, A.K. Datye, M. Neurock, R.J. Davis, J.A. Dumesic, *Journal of the American Chemical Society* 133 (2011) 12675–12689.
- [52] R.M. West, Z.Y. Liu, M. Peter, J.A. Dumesic, *ChemSusChem* 1 (2008) 417–424.
- [53] R. Johnson, *Miller & Freund's Probability and Statistics for Engineers*, 7th ed., Prentice Hall, New Jersey, 2004.
- [54] J.E. Sutton, D.G. Vlachos, *Journal of Catalysis* 297 (2013) 202–216.
- [55] N.P. Cheremisinoff, *Industrial Solvents Handbook*, 2nd ed., Marcel Dekker Inc., New York, 2003.
- [56] E.L. Kunkes, E.I. Gurbuz, J.A. Dumesic, *Journal of Catalysis* 266 (2009) 236–249.
- [57] P.A. Zapata, J. Faria, M.P. Ruiz, D.E. Resasco, *Topics in Catalysis* 55 (2012) 38–52.
- [58] A.J. Leo, D. Hoekman, *Perspectives in Drug Discovery and Design* 18 (2000) 19–38.
- [59] Biobyte Corporation, <http://biobyte.com/> (accessed 23.08.12).
- [60] J. Sangster, *Journal of Physical Chemistry Reference Data* 18 (1989) 1111–1227.
- [61] LOGKOW Database, <http://logkow.cisti.nrc.ca/logkow/index.jsp> (accessed 23.08.12).
- [62] ChemBioDraw, <http://www.cambridgesoft.com/> (accessed 23.08.12).
- [63] P. Ghosh, K.J. Hickey, S.B. Jaffe, *Industrial & Engineering Chemistry Research* 45 (2006) 337–345.
- [64] P. Ghosh, S.B. Jaffe, *Industrial & Engineering Chemistry Research* 45 (2006) 346–351.
- [65] D.A. Saldana, L. Starck, P. Mougin, B. Rousseau, L. Pidol, N. Jeuland, B. Creton, *Energy & Fuels* 25 (2011) 3900–3908.
- [66] A.R. Katritzky, M. Kuanar, S. Slavov, C.D. Hall, M. Karelson, I. Kahn, D.A. Dobchev, *Chemical Reviews* 110 (2010) 5714–5789.
- [67] H. Knozinger, *Angewandte Chemie International Edition* 7 (1968) 791–805.
- [68] H. Knozinger, H. Buhl, K. Kochloefl, *Journal of Catalysis* 24 (1972) 57–68.
- [69] C.A. Gaertner, J.C. Serrano-Ruiz, D.J. Braden, J.A. Dumesic, *Industrial & Engineering Chemistry Research* 49 (2010) 6027–6033.
- [70] A. Beretta, E. Tronconi, P. Forzatti, I. Pasquon, E. Micheli, L. Tagliabue, G.B. Antonelli, *Industrial & Engineering Chemistry Research* 35 (1996) 2144–2153.
- [71] M.H. Abraham, R.E. Smith, R. Luchtefeld, A.J. Boorem, R. Luo, W.E. Acree Jr., *Journal of Pharmaceutical Sciences* 99 (2010) 1510–1515.
- [72] C. Mintz, M. Clark, W.E. Acree Jr., M.H. Abraham, *Journal of Chemical Information & Modeling* 47 (2007) 115–121.
- [73] A. Jalan, R.W. Ashcraft, R.H. West, W.H. Green, *Annual Reports on the Progress of Chemistry Section C* 106 (2010) 211–258.